Zachary Fruhling
Writing for 3/15/07

Larry Shapiro and Elliott Sober have argued that Jaegwon Kim's Causal Exclusion Problem rests on a conceptual impossibility. Their critique is aimed at Kim's premise that mental properties can only be real if they have causal powers that their supervenience bases do not have. I will argue here that this critique is aimed to narrowly to be a successful argument against Kim's conclusions. I will argue that Kim's conclusions do not rest as heavily on the above premise as Shapiro and Sober have claimed and that in light of Kim's other reasons, Shapiro and Sober have failed argue adequately against Kim.

Shapiro and Sober hold a manipulationist account of causation based upon the work of James Woodward. According to this view, the only viable way to test for causal relations is by manipulating the supposed cause while holding fixed the background conditions and checking for a corresponding change in the cause's supposed effect. Shapiro and Sober apply this view of causation to Kim's work and argue that Kim's demand for novelty in the causal powers of a mental property above the powers of that mental property's supervenience base is a conceptual impossibility. This is impossible, Shapiro and Sober argue, because it is impossible to manipulate a supervening mental property while holding fixed the mental property's supervenience base. The strict necessity from supervenience base to

supervening property prohibits it.

It is worthwhile to spend some time reviewing the various steps of reasoning in Kim's Supervenience Argument (in which he lays out the Causal Exclusion Problem). Kim provides several reasons for thinking that a mental property's supervenience base is to be given the full causal role in bringing about an effect and that the mental property is epiphenomenal with respect both to lower level physical effects and to any mental properties that might supervene on those physical effects.

In "The Many Problems of Causal Exclusion" Kim presents what he calls the "Supervenience Argument" against Nonreductive mental causation, which starts with the assumption that the supervenience relation holds between mental and physical properties. Kim considers a case of mental-to-mental causation, in which one instance of a mental property M is supposed to be the cause of another mental property M*. According to the supervenience relation, the mental property M* has a physical supervenience base P*. The question arises: how does the mental property M* get instantiated? Kim argues that there are two competing answers to this question. The mental property M and the physical property P* appear to be competing causes in the instantiation of M*. Given that P* is necessarily sufficient for the instantiation of M*, Kim argues that the only way to reconcile the two competing claims in the production of

M* is to accept that M caused M* by causing P*. He goes far enough to offer a general principle on this insight: “To cause a supervenient property to be instantiated, you must cause its base property…to be instantiated.” (Kim in Chalmers, 176)

According to the above general principle, it seems according to Kim that mental-to-mental causation implies mental-to-physical causation. At this point in the argument, Kim notes that the mental property M must also necessarily have a physical supervenience base P according to the supervenience thesis. Kim repeats a similar line of reasoning as above, comparing the causal powers of M and the causal powers of P with respect to the production of the physical effect P*. Kim argues that on two major accounts of causation, nomological sufficiency and counterfactual causation, P has a strong claim as a cause of P*. Given that P has just as much a claim to be the cause of P* as M, as Kim eloquently states, “we are blessed with an overabundance of causes.” (Kim in Chalmers, 176). He considers the possibility that M is an intermediary cause between P and P* and rejects that possibility for the reasons that the relation between P and M is not best thought of as causal, given that in general causes precede there effects while the instantiation of a supervening property is an instantaneous occurrence.

Kim argues that if M cannot be an intermediate causal stage between P and P*, then there are only two other options remaining

that would allow M to continue to be held as a cause of P*: M and P taken together could count as a single sufficient cause of P*, or the causal powers of M and P in producing P* could be a case of causal overdetermination by two sufficient causes. Kim argues that there are two reasons for rejecting the first hypothesis that M and P together are a single sufficient cause of P*. First, he states as argued above that P alone is sufficient for P* without M, so adding M to P does not seem to add anything to the causal powers of P alone with respect to P* (this is the claim that seems best to approximate the premise Shapiro and Sober are critiquing). Second, this approach violates the principle of causal closure of the physical domain, which is that physical effects are fully explainable in terms of physical causes.

The second hypothesis, that M and P are overdetermining causes of P*, likewise has two marks against it. As above, if P is causally sufficient for M in every supposed case of mental causation, then mental properties appear to be dispensable in all cases (I assume here that Kim means that mental causes would in all cases fail to have any unique explanatory power). Second, the hypothesis that M is causally sufficient for P is also a violation of physical causal closure if one considers a possible world in which the physical cause P does not occur but the mental caused M does occur. Not only would this be a violation of causal closure, but it would be a return to a variation on

Cartesian interactionist dualism that many find untenable.

In light of the above considerations, Kim argues that there is only one plausible way out of the Causal Exclusion Problem. He argues that the best way to conceive of the situation is to posit P as the single sufficient cause of P*. P and P* each instantiate a supervening mental property, M and M*, respectively. Any supposed causal relations between M and M* or between M and P* are merely apparent causal relations that arise from the single genuine causal relation between P and P*. This proposal avoids the Causal Exclusion Problem altogether while retaining the law-like regularity of mental properties with respect to any effects that ensue.

In light of this reconstruction of Kim's argument, it seems that Shapiro and Sober have presented a false picture of the structure of Kim's reasoning. In this version of the Causal Exclusion Problem (there are admittedly variations in the ways that Kim lays out the problem in his various works), Kim does not begin with any grand claim as to the necessity of mental properties having irreducible causal powers as he does in other versions of the problem. Shapiro and Sober are really attacking Kim's reliance upon Alexander's Dictum, which is the claim that to be real is to have unique causal powers. But in the Supervenience Argument above, as I have reconstructed it, Kim does not appear to invoke Alexander's Dictum. The virtue of relying upon this particular version of the Causal

Exclusion Problem as expressed in the Supervenience argument is that this argument relies on very few problematic assumptions. One assumption is that the supervenience relation holds between mental and physical properties. A second assumption is the causal closure of the physical domain. These assumptions are both relatively unproblematic and widely accepted, and they are likely to be accepted by Shapiro and Sober as well. There is only one point at which Kim makes a claim that approximates the premise that Shapiro and Sober claim to be attacking. It occurs in Kim's critique of the hypothesis that M and P could be a single sufficient cause of P*. Kim claims that one should not entertain this hypothesis because taking P together with M would not contribute anything extra to the production of P*. Given the causal closure of the physical domain and P's sufficiency with respect to P*, Kim seems correct here without relying heavily on Alexander's Dictum. He is merely claiming that there is no positive reason to entertain the hypothesis so long as P remains causally sufficient for P* on its own. Moreover, it does not seem likely based upon Shapiro and Sober's article that they themselves are likely to entertain the notion that P and M would together constitute a single sufficient cause for P*.

Again, though: far from being a crucial step in his reasoning, this claim by Kim is being used in a very specific way to ward off a very specific hypothesis that seems to me to be the least likely of the

various hypotheses he considers to be taken seriously. It is far more likely to me that one would seriously entertain the overdetermination hypothesis. His positive arguments against the overdetermination hypothesis aside, the important realization for the present purpose is that nowhere in his response to the overdetermination claim does he invoke the principle that Shapiro and Sober are attacking. Kim may very well believe the Alexander's Dictum thesis to be true, and it may even be lurking in the background providing the *motivation* for Kim's argument in the first place. But the fact that this claim motivates the argument does not ipso facto make the claim a crucial premise in the argument. As I have shown above, Kim's argument actually relies very little, if at all, on a strong version of Alexander's Dictum. The Causal Exclusion Problem, if properly set up, relies on a minimum of relatively unproblematic assumptions, which is one of the hallmarks of any good philosophical arguments. This is not to say that there are not other problems with the argument than the problem raised by Shapiro and Sober, but Shapiro and Sober have provided an argument that is poorly aimed at Kim's actual reasoning. Instead they offer a type of straw man argument that does not correctly portray Kim's weak use of Alexander's Dictum, making that claim out to be a key assumption when it is nothing of the sort.